AD_____

Award Number: DAMD17-98-1-8239

TITLE: Rolling Circle Transcription of Ribozymes Targeted to ras and mdr-1

PRINCIPAL INVESTIGATOR: Eric T. Kool, Ph.D.

CONTRACTING ORGANIZATION: University of Rochester
Rochester, ~~NY~~ 14627-0140

REPORT DATE: September 2000

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 2000	Annual (1 Sep 99 - 31 Aug 00)

4. TITLE AND SUBTITLE
Rolling Circle Transcription of Ribozymes Targeted to ras and mdr-1

5. FUNDING NUMBERS
DAMD17-98-1-8239

6. AUTHOR(S)
Eric T. Kool, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Rochester ~~and Stanford University, Stanford CA 94305~~
Rochester, New York 14627-0140

E-MAIL:
kool@leland.stanford.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

We report new progress on the use of the rolling circle transcription strategy for production of ribozymes in vitro and in living cells. Optimized circular DNA vectors have been selected from randomized libraries and the selected sequences were found to be transcribed significantly more strongly than the library as a whole by two different RNA polymerases. Optimized rolling circle vectors were then constructed to encode a ribozyme targeted to a drug resistance gene in a bacterium. The vector was then shown to lower the activity of this gene in the living bacterial cells. A similar strategy may one day be useful for inhibiting genes (such as breast cancer-associated genes) in human cells.

14. SUBJECT TERMS
Breast Cancer

15. NUMBER OF PAGES
42

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

TABLE OF CONTENTS

Table of Contents

BODY OF REPORT

The information in this document is proprietary to the University of Rochester and to Stanford University.

BCRP / IDEA grants program
GRANT # DAMD17-98-1-8239

TITLE: **Rolling Circle Transcription of Ribozymes Targeted to *ras* and mdr-1**

P.I.: KOOL, ERIC T.
Professor of Chemistry, Department of Chemistry, Stanford University, Stanford, CA 94305-5080; Tel (650) 724-4741; kool@stanford.edu

INTRODUCTION

This IDEA project focuses on a new method for generating biologically active ribozymes (1-11). This method is termed "rolling circle transcription", and involves unusual circular single-stranded DNA templates. We have previously discovered that synthetic DNA circles as small as ~30 nucleotides in length could be transcribed efficiently by RNA polymerases, despite their lack of promoter sequences. When we encode ribozymes in these circles, as well as their self-cleavage sequences, then this produces a long string of ribozymes that self-cleaves until virtually the only product is a ribozyme (in amplified quantities) that has the same length as the circle (12-15).

We proposed the interesting possibility that such circular DNAs might one day be used in human cells to encode biologically active ribozymes. These ribozymes might target disease-related RNAs for destruction, and thus the circular DNAs could have a biological effect. In breast cancer, two mRNA targets that are likely to be important are H-ras and mdr-1. H-ras is an oncogene that is commonly mutated and overexpressed in breast cancer; it is possible that downregulation of this mutated gene would affect cancer growth in a favorable way. The gene mdr-1 is very commonly overexpressed in malignancies that Have undergone anticancer therapy; this leads to drug resistance in further treatment. Thus, downregulation of mdr-1 RNA might well lead to an enhanced ability to treat breast cancer with standard drugs.

Before our synthetic circular DNAs can be used as vectors in breast cancer therapy, quite a number of questions need to be answered, and some of these are the subject of this ongoing project. We need to measure the RNA-cleaving activity of ribozymes produced by rolling circle transcription. Are they as efficient as standard ribozymes? We also need to find sequences of DNA circles that are most efficiently

transcribed. Can transcription be made more efficient than we have already observed? We need to find sequences that lend optimal activity and stability in cellular media. Finally, we would like to optimize transcription by cellular polymerases. Can these more complex polymerases utilize such small circles as templates?

Below is an outline of the proposed statement of work for this project:

Task 1. Compare cleavage activity and nuclease stability of self-processed vs. conventional ribozymes

Task 2. Optimize in vitro transcription and investigate incorporation of modified ribonucleotides

Task 3. Evaluate the most active and stable RNAs in cell culture

Task 4. Carry out in vitro selection to optimize transcription of such vectors by human RNA polymerases

In the second year of this project we have made significant progress on all proposed tasks, with strongest progress in tasks 2-4, having completed task 1 as planned. Below are described details of our work on the remaining three tasks.

TASK 1: Cleavage activity of self-processed ribozymes:

--completed in year 1--

TASK 2: Optimizing in vitro transcription:

As mentioned in the introduction, before this approach can be generally applied we need to find ways to enhance transcription and make it predictable and reliable regardless of the particular ribozyme encoded in the vector. Over the past year we have completed an optimization of transcription by two different classes of polymerases, one a viral RNA polymerase (T7 RNA polymerase), and one a bacterial polymerase (E. coli RNA polymerase).

A 103-nt circular DNA was prepared, encoding the 63-nt mdr-1 ribozyme as well as a 40 nt randomized domain. The method used to construct the molecule was as done before, using two shorter linear DNAs and ligating them in two steps. This yields a circular DNA library of ca. 10^{13} different sequences.

This library was then used in an in vitro selection scheme, in which we select for sequences that are transcribed and yield monomer ribozyme RNAs. The scheme briefly works as follows: the library is transcribed and the RNAs are allowed to self-process. They are separated on a gel next to a size marker. The monomer-length RNAs are excised from the gel. They are amplified by RT-PCR and then the DNAs are re-cyclized. This new set of circular DNAs has thus been enriched in sequences that are especially well transcribed. Moreover, the corresponding RNAs had to retain high self-cleavage activity to appear as the

monomer band on the gel. Additional rounds lead to further enrichment. At the end of the selection, the PCR fragments are cloned and sequenced. This yields optimized circle sequences, all of which encode the mdr-1 ribozyme.

We carried out 10-16 rounds of selection in this fashion, using two RNA polymerases in separate experiments: T7 RNAP, and E. coli RNAP. Cloning resulted in the identification of "winner" sequences for both enzymes. We then prepared new DNA circles having single sequences based on the various classes of winner sequences identified in the selection. We transcribed them and evaluated the efficiency relative to the original library and relative to the original MDR63 circular DNA.

Results showed in some cases greatly enhanced activity. In the E. coli RNA polymerase case, we identified a circle sequence that is transcribed up to 80 times more efficiently than the control. In the T7 case we found a winner that is 5-fold more active.

Subsequent experiments (see below) showed that the selected DNA domains that arose from the randomized segment could be transplanted to a new DNA circle encoding a new ribozyme. We constructed a new circle encoding a ribozyme targeted to a bacterial drug resistance gene, marA. This circle included a selected domain that was a "winner" in the E. coli RNAP selection experiments. Transcription of this new circle showed that it was in fact transcribed even more strongly than the mdr-1 selected 103mer circle.

TASK 3: Evaluate most active RNAs in cell culture:

Our long-term goal is to engender transcription in human cells using rolling circle vectors. However, because of the relatively slow progress made in finding winning sequences for transcription in HeLa extract (see below), we opted to attempt cellular studies in a bacterial system to try to achieve proof-of-principle.

As this work is still being completed, we will only briefly describe it here; a full description will be given in the final report for this project.

A 103-nt circular DNA was constructed. 40 nt were taken directly from the winner sequence from the E. coli in vitro selection experiments above. The remaining 63nt encoded a hammerhead ribozyme targeting the marA antibiotic resistance gene in E. coli. To test this, a plasmid was constructed encoding a segment of the marA gene upstream from a CAT reporter gene. Cleavage of the marA segment is therefore expected to result in loss of CAT activity in the bacteria.

We used heat shock to allow the circular DNA to permeate the cells. After a short incubation we measured CAT activity. Results showed that the marA vector resulted in loss of up to 80% of CAT activity, while the same circle encoding an inactivated ribozyme had no effect.

Thus we conclude that small synthetic "rolling circle" DNAs can in fact encode active ribozymes in bacterial cells.

MATERIALS AND METHODS

Preparation of Oligonucleotides and Library Circular DNA. DNA oligonucleotides were synthesized on solid supports using the phosphoramidite method on an Applied Biosystems model 392 DNA/RNA synthesizer. Oligodeoxyribonucleotides were deprotected by treatment with concentrated 25% ammonia at 55°C for 8 h. After deblocking, DNA oligonucleotides were purified by electrophoresis on polyacrylamide denaturing gels (PAGE). After elution from the gels, the oligonucleotides were desalted again with C18 Sep-Pak cartridges. Single-strand concentrations of purified DNA oligonucleotides were determined by measuring the absorbance at 260 nm at high temperature. Single-strand extinction coefficients were calculated from mononucleotide and dinucleotide data with a nearest-neighbor approximation.

An initial circular ssDNA library containing 63 nt fixed sequence and 40 nt randomed sequence was generated by sequential enzymatic ligations of 5'-phosphorylated 56 nt and 47 nt oligonucleotides using T4 DNA ligase (New England Bio Labs, Inc.) and 16 nt splint oligonucleotides as previously described. The 5'-phosphorylated sequences were:, 5'-pTTC GTC TG-N_{40}-TCT TTC AG -3', 5'- TTT CGT CCT CAC GGA CTC ATC AGA ATG GCA ACA CAT TGA CTG AGG AG-3'.

In Vitro Selection. Condition for initial rolling circle transcription reaction were: 1 μM circular DNA, 2 units *E. coli* RNA polymerase (RNAP) holoenzyme, 0.5 mM ATP, CTP, GTP, 60 μM UTP, 0.30 μCi of _-[^{32}P] UTP in 25 mM Tris-HCl (pH 8.1) buffer containing 20 mM NaCl, 15 mM $MgCl_2$, 0.4 mM spermine-HCl, 100 μg/mL acetylated bovine serum albumin, 10 mM DTT, and 12.5 units/mL RNase inhibitor, in a total reaction volume of 15 μL. After 1.5 h incubation at 37°C, the reaction was terminated by adding an equal volume of stop solution (30 mM Na_2EDTA, 8 M urea, 0.02% bromophenol blue, and 0.02% xylene cyanol). Self-processed 103 nt product RNA was purified by 10% denaturing PAGE that run at 4°C.

After elution from the gels, the selected 103 nt ssRNAs were reverse transcribed using reverse transcriptase (Invitrogen) and 5'-phosphorylated 18nt primer (pGAC TGA GGA GTT CGT CTG) for 1h at 42°C. After buffer exchange with Bio-spin column (Bio-Lad), the cDNA products were PCR-amplified *Taq* polymerase by 15 cycles (temperature cycle: 94°C, 1min; 55°C, 1min; 72°C, 1min) in the presence of 100 pmols each primers, 5'-pGAC TGA GGA GTT CGT CTG-3' and 5'-biotin-AAT GTG TTG CCA TTC TGA-3'. The PCR products were extracted with phenol/$CHCH_3$, ethanol precipitated, and blunted with T4 DNA polymerase (LifeTechnologies, Inc.).

The PCR products with blunted ends were immobilized on magnetic beads with streptavidin (Dynal) in the presence of 80 μM binding buffer (1M NaCl, 10mM Tris-HCl (pH7.5), 1mM EDTA), rinsed with two 80 μL volumes of the binding-buffer, and eluted with 70 μL volume of 0.15 N NaOH to recover the non-biotinylated ssDNA. Then the recovered solution was exchanged with spin-column.

After quantifying the non-biotinylated ssDNA with 5'-phosphate by measuring the absorbance at 260 nm, 1 μM non-biotinylated ssDNA with 5'-phosphat was mixed with 2 μM 24 nt splint DNA (5'-AAC TCC TCA GTC AAT GTG TTG CCA-3') in the presence of 10 mM Tris-HCl (pH 7.5) and 10 mM $MgCl_2$, annealed with thermal cycler (rate: 1°C/min) from 90°C to 25°C, and incubated with 0.3 units T4 DNA ligase, 0.1 mM DTT, and 10 mM ATP at 25°C for 12h in a final total reaction volume of 15 μL. The circular ssDNA, unreacted ssDNA, and splint ssDNA were ethanol-precipitated and used as templates to begin the next round of in vitro selection without purification. The rolling circle transcription reaction for next round was done under the condition as indicated above except template DNA concentration.

Cloning and Sequencing. The polymerase chain reaction products of the fifteenth round pool were ligated into a TA cloning vector (Invitrogen,) and cloned into *E. coli* TOP10F' (Invitrogen). Plasmid DNAs were isolated and sequenced using BigDye terminator cycle sequencing kit (PE Applied Biosystems).

Synthesis of Selected Circular ssDNAs and Measurement of their Activities. All selected and redesigned circular ssDNAs were made by sequential enzymatic ligations as indicated above.

All rolling circle transcription reactions were done under condition as indicated above using *E. coli* RNAP with 0.2 μM the circular ssDNA templates. Reactions were incubated at 37°C, and were stopped by the addition of one volume of stop solution. Gel analysis was on a 10% PAGE run at 4°C. The transcribed and self-processed RNAs were quantified with a radioanalytical scanner (Molecular Dynamics Storm 860).

Preparation of DNA Plasmid Containing Target MAR A Gene and CAT Gene. We created DNA fragment containing MAR A gene and CAT gene by two PCR steps with synthetic primers. At first step, the CAT gene fragment was amplifled form pKK232-8 plasmid (Amersham Pharmacia Biotech) by PCR with 29 nt and 27nt primers. . Primer sequence were: 5'-AGG TCG ACT ATG GAG AAA AAA ATC ACT GG-3' and 5'-GGT ACC CAA AAG GCC ATC CGT CAG GAT-3'. After purification from agarose gel, the PCR product was amplifled by PCR with 81 nt and 29 nt primers. One 81 mer primer had 51 MAR A gene and Hind III site. Another 29 nt primer has KpnI site. Primer sequence were: 5'-CCC AAG CTT GTC ACT GGA GAA AGT GTC AGA GCG TTC GGG TTA CTC CAA ATG GCA CCT GCA AAT GGA GAA AAA AAT CAC TGG and 5'-GGG GTA CCC AAA AGG CCA TCC GTC AGG AT. As result, 51mer segment of the MAR A gene was upstream of the CAT gene. After purification from agarose gel, this MAR A-CAT fragment (1300 nt) was ligated into pUC19 (Amersham Pharmacia Biotech) at the HindIII site and KpnI site with T4 DNA ligase. The ligated plasmides were cloned into INV_F' *E. coli* (Invitrogen) and picked up by blue-white color screening.

CAT Assay. INV_F' *E. coli* stain (Invitrogen) was used as cell. After cells were transfected with the CAT vector, cells were cultured in LB containing 50 μg/mL ampicillin until the absorbance at 600 nm of 0.1. Then we transfered 1 μL cell solution with 0.1 absorbance into 100μL LB solutiosn with the circular DNA and incubated until the absorbance at 600 nm became 0.1-0.2. We also did heat-shock (42 °C) at every 1 h. The CAT activities were measured with [^{14}C] chloramphenicol (100μCi/ml) and a CAT enzyme assay system (Promega). In order to normalize the efficiency of protein extract by reference to ß-galactosidase activity, cells were cotransfected with pSV-ß-Galactosidase Control Vector (Promrga) and then the chemiluminescent signal due to ß-galactosidase was determined with ß-galactosidase enzyme system (Promega).

Check the cleavage of mRNA by the ribozyme using RT-PCR. After cell was incubated with or without the circular DNA until the absorbance at 600 nm became 0.1-0.2, the total RNAs were isolated with SV Total RNA Isolation System (Promega). The isolated RNAs were reverse transcribed using ThermoScript RT-PCR system (LifeTechnologies,Inc.) and 22 nt primer (5'-GTA TAT CCA GTG ATT TTT TTC T-3') for 1h at 65°C. Then the cDNA products were PCR-amplified by 25 cycles (temperature cycle: 94°C, 1min; 55°C, 1min; 72°C, 1min) in the presence of two different primer sets. One is : 5'-GTA TAT CCA GTG ATT TTT TTC T-3' and 5'ATG ACC ATG ATT ACG CC -3'. Other is : 5'-GTA TAT CCA GTG ATT TTT TTC T-3' and 5'AGA GCG TTC GGG TTA CTC CA -3'. The PCR products were quantified with NIH Image.

TASK 4: In vitro selection with human RNA polymerases:

Our approach for finding circular DNA sequences that are strongly transcribed by human RNA polymerases is the same as was successful with bacterial enzymes. In fact, the same original library was the one that has been used in the human selection. Our source of active human RNA polymerases has been a commercially available HeLa cell nuclear extract. However, although the bacterial enzyme selection showed very good progress, the HeLa case has not as yet. We believe this is due largely to the weak transcriptional activity present in the extract. Nonetheless we are continuing our rounds of selection in the hope that there are rare winner sequences that will emerge at later rounds. If this is not successful, then we will repeat the experiments with a new circular DNA library having a larger randomized domain.

KEY RESEARCH ACCOMPLISHMENTS

• We have identified new classes of single-stranded DNA "promoters" that are highy active in transcription by viral and bacterial enzymes.

• We have established that "rolling circle vectors" can show high activity in delivering biologically active ribozymes into bacterial cells.

REPORTABLE OUTCOMES

One paper has appeared in print since the last report:
Ohmichi, T. and Kool, E.T. (2000) *Nucleic Acids Res.*, **28**, 776-83.

One paper has been accepted for publication and is now in press:
Eric T. Kool, Modified DNAs as substrates for polymerases, *Curr. Opin. Chem. Biol.* **2000**, in press.

CONCLUSIONS

(1) In vitro selection methodology has allowed us to find specific DNA circle sequences that are even more highly efficient tat average sequences in transcription, at least for bacterial and viral polymerases.
(2) We have shown that "rolling circle vectors" can show high activity in delivering biologically active ribozymes into bacterial cells, cleaving a segment of a drug resistance gene and downregulating gene activity in the living cells.

REFERENCES

1. Buzayan,J.M., Gerlach,W.L., and Bruening,G. (1986) *Nature*, **323**, 349-353.

2. Hutchins,C.J., Rathjen,P.D., Forster,A.C., and Symons,R.H. (1986) *Nucleic Acids Res.*, **14**, 3627-3640.

3. Prody,G.A., Bakos,J.T., Buzayan,J.M., Schneider,I.R., and Bruening,G. (1986) *Science*, **231**, 1577-1580.

4. Forster,A.C. and Symons,R.H (1987) *Cell*, **49**, 211-220.

5. Uhlenbeck,O.C. (1987) *Nature*, **328**, 596-600.

6. Haseloff,J. and Gerlach,W.L. (1988) *Nature*, **334**, 585-591.

7. Zhou,D.-M. and Taira,K. (1998) *Chem. Rev.*, **98**, 991-1026.

8. Eckstein,F. and Lilley,D.M.J., eds. (1996) *RNA Catalysis*: *Nucleic Acids and Molecular Biology* , Springer Verlag, Berlin, Vol. 10.

9. Chen,H., Ferbeyre,G. and Cedergren,R. (1997) *Nature Biotechnol.*, **15**, 273-277.

10. Kuwabara,T., Warashina,M., Nakayama,A., Ohkawa,J. and Taira,K. (1999) *Proc. Natl. Acad. Sci. USA*, **96**, 1886-1891.

11. Hertel,K.J., Herschlag,D., and Uhlenbeck,O.C. (1996) *EMBO J.*, **15**, 3751-3757.

12. Daubendiek, S.L., Ryan,K., and Kool,E.T. (1995) *J. Am. Chem. Soc.*, **117**, 7818-7819.

13. Daubendiek, S.L. and Kool,E.T. (1997) *Nature Biotechnol.*, **15**, 273-277.

14. Diegelman, A.M. and Kool,E.T. (1998) *Nucleic Acids Res.*, **26**, 3235-3241.

15. Diegelman,A.M. and Kool,E.T. (1999) *Chem. & Biol.*, **9**, 569-576.

16. Usman,N., Ogilvie,K.K., Jiang,M.-V., and Cedergren,R. (1987) *J. Am. Chem. Soc.*, **109**, 7845-7854.

APPENDIX

Copies of one reprint, and one paper in press:

Ohmichi, T. and Kool, E.T. (2000) *Nucleic Acids Res.*, **28**, 776-83.

Eric T. Kool, Modified DNAs as substrates for polymerases, *Curr. Opin. Chem. Biol.* **2000**, in press.

776–783 Nucleic Acids Research, 2000, Vol. 28, No. 3

The virtues of self-binding: high sequence specificity for RNA cleavage by self-processed hammerhead ribozymes

Tatsuo Ohmichi and Eric T. Kool*

Department of Chemistry, University of Rochester, Rochester, NY 14627, USA

Received September 15, 1999; Revised and Accepted December 3, 1999

ABSTRACT

Naturally occurring hammerhead ribozymes are produced by rolling circle replication followed by self-cleavage. This results in monomer-length catalytic RNAs which have self-complementary sequences that can occupy their *trans*-binding domains and potentially block their ability to cleave other RNA strands. Here we show, using small self-processed ribozymes, that this self-binding does not necessarily inhibit *trans*-cleavage and can result in greatly elevated discrimination against mismatches. We utilized a designed 63 nt circular DNA to encode the synthesis of a self-processed ribozyme, MDR63. Rolling circle transcription followed by self-processing produced the desired 63 nt ribozyme, which potentially can bind mdr-1 RNA with 9+9 nt of complementarity or bind itself with 4+5 nt of self-complementarity by folding back its ends to form hairpins. Kinetics of *trans*-cleavage of short complementary and mismatched RNAs were measured under multiple turnover conditions, in comparison to a standard 40 nt ribozyme (MDR40) that lacks the self-complementary ends. The results show that MDR63 cleaves an mdr-1 RNA target with a k_{cat}/K_m almost the same as MDR40, but with discrimination against mismatches up to 20 times greater. Based on folding predictions, a second self-processed ribozyme (UG63) having a single point mutation was synthesized; this displays even higher specificity (up to 100-fold) against mismatches. The results suggest that self-binding ends may be generally useful for increasing sequence specificity of ribozymes.

INTRODUCTION

The hammerhead ribozyme was originally identified as the catalytic RNA motif responsible for self-processing of plant infecting viroid RNAs (1–4). The replication pathway for these infectious RNAs involves rolling circle replication to produce a multimeric copy and requires a self-cleavage step to produce monomer-length viroid RNAs in amplified quantities. Plant viroids are typically 200–400 nt in length, but the catalytic function can be made considerably smaller. Minimal hammerhead RNA motifs have been identified that can cleave other RNAs with multiple turnovers (5–7) and such agents are under investigation as therapeutic agents for specific gene inhibition (8–10). It has been observed, however, that the specificity of RNA cleavage by hammerhead ribozymes is low, especially if the ribozyme/target RNA recognition helices are long (11). When these helices are too short, however, cleavage site specificity is likely to be adversely affected in the presence of complex nucleic acid sequences. Because cleavage of mismatched targets is undesirable in therapeutic and diagnostic strategies, there would be considerable value in finding ways to increase sequence specificity of ribozyme cleavage.

We have previously shown that small synthetic circular single-stranded DNAs can behave as efficient templates for RNA polymerases, despite their lack of promoter sequences (12). The resulting rolling circle transcription produces long multimeric RNAs in a mechanism that mimics the first step of viroid RNA replication. Further, when ribozyme RNAs and their cleavage substrates are encoded in such circular vectors, the repeating RNAs undergo self-processing, yielding monomer-length ribozyme RNAs as the chief products and in amplified amounts. This biomimetic strategy has been successfully used in the synthesis of hammerhead, hairpin and hepatitis delta ribozyme motifs (13–15).

As with viroid monomer RNAs, however, the monomer ribozymes produced from such a self-processing mechanism necessarily contain not only the minimal catalytic RNA, but also self-complementary RNAs which are the remnants of self-cleavage. These sequences are located at the ends of the minimal ribozyme and have the potential to fold back and occupy the domains of the ribozyme that are necessary for binding to another RNA before cleaving it. Thus there is the possibility that this self-binding might interfere with *trans*-cleaving ability. For this reason, internal self-complementarity might well be expected to strongly affect the *trans*-cleaving properties of ribozymes. Importantly, although all hammerhead ribozymes in their natural form do possess such complementarity, the effect on *trans*-cleaving efficiency and specificity has apparently not been investigated.

Here we show that self-complementary ends resulting from self-processing do not necessarily inhibit *trans*-cleaving activity of a hammerhead ribozyme. Moreover, we find that

*To whom correspondence should be addressed at present address: Department of Chemistry, Stanford University, Stanford, CA 94305-5080, USA. Tel: +1 650 724 4741; Fax: +1 650 725 0259; Email: kool@leland.stanford.edu

sequence specificity is markedly increased by this potential structure. The results suggest that rolling circle transcription coupled with self-processing may be a generally useful strategy for synthesis of ribozymes that have enhanced properties for RNA cleavage.

MATERIALS AND METHODS

Preparation of oligonucleotides and circular DNAs

DNA and RNA oligonucleotides were synthesized on solid supports using the phosphoramidite method on an Applied Biosystems model 392 DNA/RNA synthesizer. Oligodeoxyribonucleotides were deprotected by treatment with concentrated 25% ammonia at 55°C for 8 h (16). Synthesized RNAs were removed from the solid support and base blocking groups were removed by treatment with concentrated 25% ammonia in ethanol (3:1 v/v) at 55°C for 8 h. After drying in vacuum, the 2′-silyl protecting groups were removed by resuspending the pellet in 50 equiv. of 1.0 M tetrabutylammonium fluoride per equivalent of silyl and the mixtures incubated overnight at room temperature. The oligoribonucleotides were then passed through a C-18 Sep-Pak cartridge column for desalting. After deblocking, DNA and RNA oligonucleotides were purified by electrophoresis on polyacrylamide denaturing gels. After elution from the gels, the oligonucleotides were desalted again with C18 Sep-Pak cartridges. Construction of circular DNAs was done by sequential enzymatic ligations. Ligations of 5′-phosphorylated 35 and 28 nt oligonucleotides were performed sequentially using T4 DNA ligase and 24 nt splint oligonucleotides as described previously (13,14,17). Sequences were: MDR63, 5′-pGAC TGA GGA GTT CGT CTG TCT TTC AGT TTC GTC CT-3′ and 5′-pCAC GGA CTC ATC AGA ATG GCA ACA CAT T-3′; UG63, 5′-pGAC TGA GGA GTT CGT CTG TCT TTC AGT TTC GTC CT-3′ and 5′-pCAC GGA CTC ATC AGA ATG GCA ACC CAT T-3′. Single-strand concentrations of purified DNA and RNA oligonucleotides were determined by measuring the absorbance at 260 nm at high temperature. Single-strand extinction coefficients were calculated from mononucleotide and dinucleotide data with a nearest neighbor approximation (18). The MDR40 short ribozyme was synthesized by T7 RNA polymerase transcription of a synthesized DNA template and purified by 15% denaturing gel electrophoresis (19). Sequences of RNA targets of varying length were 5′-UCA GUA AAU GG-3′, 5′-UUU CAG UCA AUG GCA-3′, 5′-UGU UUC AGU CAA UGG CAA C-3′ and 5′-CUG UGU UUC AGU CAA UGG CAA CAC A-3′.

Transcription reactions

Conditions for an internally labeled rolling circle transcription reaction were: 1 µM circular DNA, 2 U *Escherichia coli* RNA polymerase (RNAP) holoenzyme or 25 U T7 RNAP, 0.5 mM ATP, CTP and GTP, 60 µM GTP, 0.30 µCi [α-^{32}P]UTP in 25 mM Tris–HCl (pH 8.1) buffer containing 20 mM NaCl, 15 mM $MgCl_2$, 0.4 mM spermine–HCl, 100 µg/ml acetylated bovine serum albumin, 10 mM DTT and 12.5 U/ml RNase inhibitor, in a total reaction volume of 15 µl. Unlabeled MDR63 and UG63 were prepared from a rolling circle transcription reaction as indicated above using *E.coli* RNAP with all four rNTPs at 0.5 mM. Reactions were incubated at 37°C for 12 h and the reaction was stopped by the addition of 1 vol of stop solution (30 mM EDTA, 8 M urea). Gel analysis was on a 10% polyacrylamide denaturing gel run at 4°C.

Sequencing of monomer ribozymes

5′-End-labeling with T4 polynucleotide kinase was done for monomer RNAs following standard procedures. RNase T1 cleavage and alkaline hydrolysis was performed on the ^{32}P-labeled monomers after they were ethanol precipitated and redissolved in water. Alkaline hydrolysis was carried out in 50 mM sodium bicarbonate (pH 9.0) buffer containing 1 mM EDTA for 10 min at 90°C. RNase T1 cleavage conditions were as follows: ^{32}P-labeled RNA and 0.064 U/µl RNase T1 (US Biochemical) in 20 mM sodium citrate (pH 3.5) buffer containing 6 M urea and 1 mM EDTA, reacted at 50°C for 5 min. All reactions were stopped by rapid cooling on dry ice prior to immediate analysis on a 15% polyacrylamide denaturing gel.

Cleavage reactions

Multiple turnover experiments were performed with RNA substrate in at least 10-fold excess over the ribozyme (20). The ^{32}P-labeled RNA substrate and 50 nM ribozyme were separately heated to 90°C for 1 min, cooled slowly and incubated at 37°C for 20 min. All cleavage reactions under the multiple turnover conditions were initiated by mixing the substrate and the ribozyme. In the case of single turnover conditions, each ribozyme and substrate was heat treated in reaction buffer separately and then allowed to reach reaction temperature. Reactions were initiated by combining various concentrations (50 nM–1.5 µM) of ribozyme and a 5 nM ^{32}P-labeled RNA substrate. A different reaction protocol was also used in which ribozyme (50 nM–1.5 µM) and ^{32}P-labeled RNA substrate (5 nM) were heat treated together in 25 mM Tris–HCl (pH 8.1) buffer containing 20 mM NaCl and 0.4 mM spermine. After equilibration to 37°C, the reaction was initiatied by the addition of $MgCl_2$ to a final concentration of 15 mM. The reactions were terminated by adding an equal volume of 100 mM Na_2EDTA, 9 M urea, 0.02% bromophenol blue and 0.02% xylene cyanol. The labeled product and substrate were separated by electrophoresis on 20% polyacrylamide denaturing gels. The RNA cleavage yields were determined by quantitation of radioactivity in the bands of labeled products and substrate with a radioanalytical scanner (Molecular Dynamics Storm 860). Initial rates corresponding to the first 20% of reaction were used to obtain rate constants and the k_{cat} and K_m values were calculated from non-linear least square fits (KaleidaGraph; Ablebeck Software). Rate constants for reaction under single turnover conditions were determined from the slopes of semi-logarithmic plots of the ^{32}P-labeled RNA substrate concentration, normalized to the final extent of cleavage, versus time.

RESULTS

Design of circular vector and *in vitro* transcription

In the case of some kinetic investigations of the mechanism of the hammerhead ribozyme, short recognition stems (e.g. 5–6 nt) are used because the relatively low binding affinity leads to high activity (21). In the case of self-processing ribozymes, however, a short recognition stem would not be sufficient for high *trans*-cleavage activity (shown below), since the

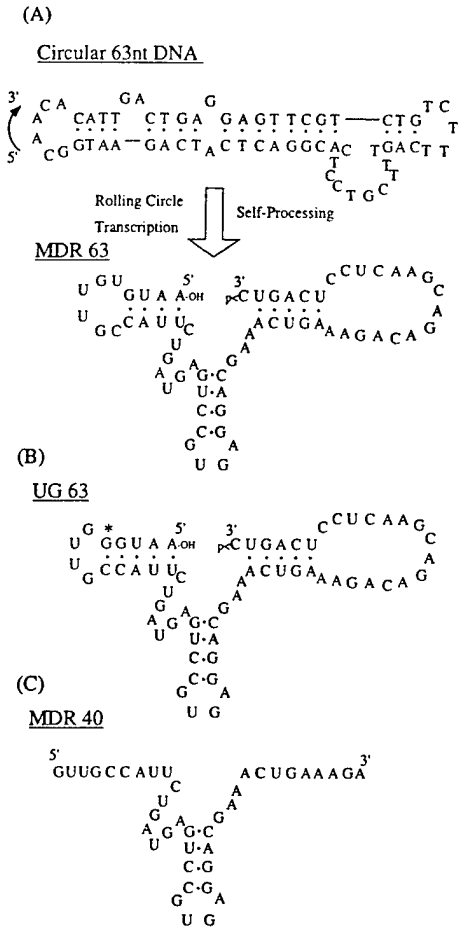


Figure 1. (**A**) Sequences and predicted secondary structures of the 63 nt circular DNA encoding MDR63 and the product, the self-processed MDR63 hammerhead ribozyme. (**B**) Sequence of the self-processed UG63 ribozyme. The single base changed from MDR63 is marked by an asterisk. (**C**) Sequence of the minimal MDR40 ribozyme, which is the same as the above two ribozymes but without the self-complementary ends.

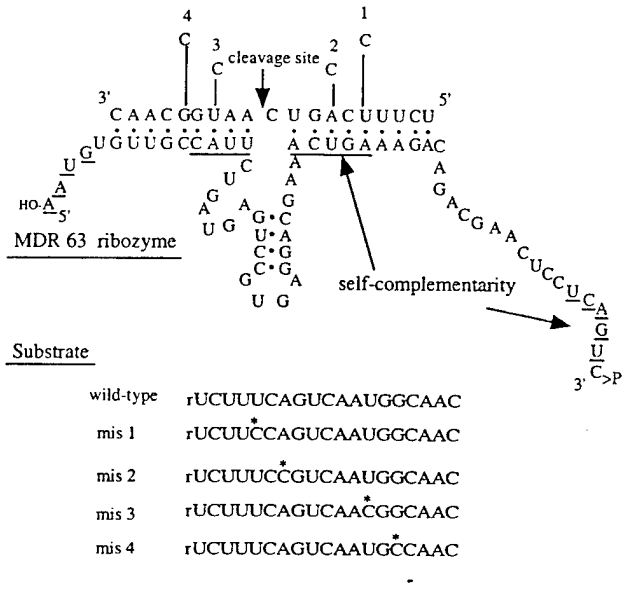


Figure 2. Illustration of the MDR63 ribozyme–substrate complex and sequence of the four mismatched substrates studied. Mismatched bases are marked by an asterisk. Self-complementary bases are underlined; any self-binding that involves these must be displaced by an incoming substrate RNA.

ribozyme must have complementarity (at least 4 or 5 nt) with itself to make the intramolecular hammerhead ribozyme structure. Thus, there must be enough binding energy between the substrate and the recognition domains of the ribozyme to overcome this self-binding. For that reason, 9 nt of recognition was designed as the stem length on each side of the cleavage domain. For a minimal hammerhead RNA designed to bind 9+9 nt of an RNA target, a total size of 39–40 nt is necessary. To generate a hammerhead RNA with this activity by self-processing, one must also include ~16–28 nt of sequence that can be cleaved by this ribozyme. We previously showed that an 83 nt DNA circle could be transcribed to yield self-processing hammerhead RNAs (13,22); in the present case we shortened the length by 20 nt, giving a 63 nt construct, MDR63 (Fig. 1). Because of concerns that too much self-complementarity might greatly limit *trans*-cleaving activity, we designed MDR63 to have 9+9 nt of complementarity for mdr-1 RNA (23), but to have less complementarity (4+5 nt) for itself. RNA folding analysis (24–27) predicts free energies (37°C) of –0.2 and –3.2 kcal/mol for the two self-complementary hairpins in MDR63 (Fig. 1A) and –11.3 and –10.3 kcal/mol for the helices formed by binding a perfectly complementary target (as in Fig. 2). Also synthesized for comparison was the minimal hammerhead MDR40, which is identical to MDR63 but lacks the self-complementary ends (Fig. 1C).

Studies were then carried out to determine whether the circular DNA encoding MDR63 was transcribed *in vitro* by *E.coli* or T7 RNA polymerase (RNAP). Figure 3A shows the products after 12 h of transcription at 37°C. Both enzymes produced virtually identical products, which appear as several bands on the polyacrylamide gel. These were presumed to be monomer, dimer and higher order multimers, similar to previous findings with transcription of circular ribozyme DNAs (13–15). RNase T1 sequencing confirmed that the fastest migrating band was the expected MDR63 linear monomer RNA (Fig. 3B). This ribozyme was then prepared in larger amounts in unlabeled form for kinetics studies.

Basic *trans*-cleavage properties of MDR63

Kinetics studies were initially carried out under multiple turnover conditions to find the *trans*-binding helix length that is optimum for highest cleavage efficiency with this self-complementary ribozyme. The effect of length was investigated with four target RNAs of increasing size: 11, 15, 19 and 25 nt (see Materials and Methods for sequences). These have 5+5, 7+7, 9+9 and 12+12 nt of complementarity for the ribozyme, respectively. Kinetics studies were carried out under multiple turnover conditions in a buffer containing 15 mM $MgCl_2$, 20 mM NaCl and 0.4 mM spermine at 37°C, conditions similar to those used for transcription of the ribozyme itself. This allowed us to examine cleavage under the conditions where self-cleavage was known to occur successfully.

Observed second order rate constants (k_{cat}/K_m values) were 8.0×10^3, 6.0×10^4, 2.6×10^5 and 1.7×10^4 M^{-1} min^{-1} with the 11mer, 15mer, 19mer and 25mer substrates, respectively. Thus, the experiments showed that 9+9 nt of complementarity gave the highest cleavage efficiency under these conditions.

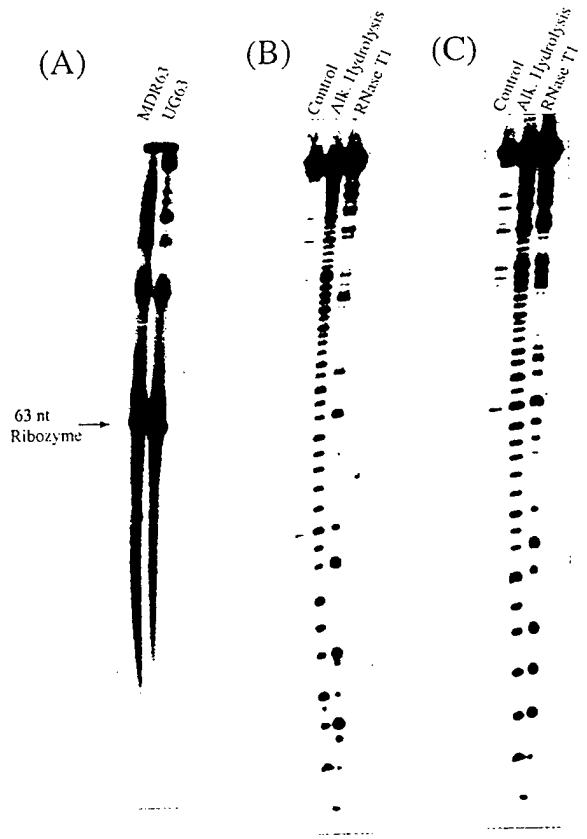


Figure 3. Transcription and characterization of self-processed ribozymes. (**A**) Autoradiogram of denaturing 10% polyacrylamide gel showing *in vitro* transcription of the 63 nt circular DNAs encoding MDR63 and UG63 by *E.coli* RNAP (after 12 h) and showing the monomer bands that arise after self-processing. (**B**) Autoradiogram of denaturing 15% polyacrylamide gel showing RNase T1 sequencing of the monomeric MDR63 hammerhead ribozyme. (**C**) Sequencing of the monomeric UG63 hammerhead ribozyme.

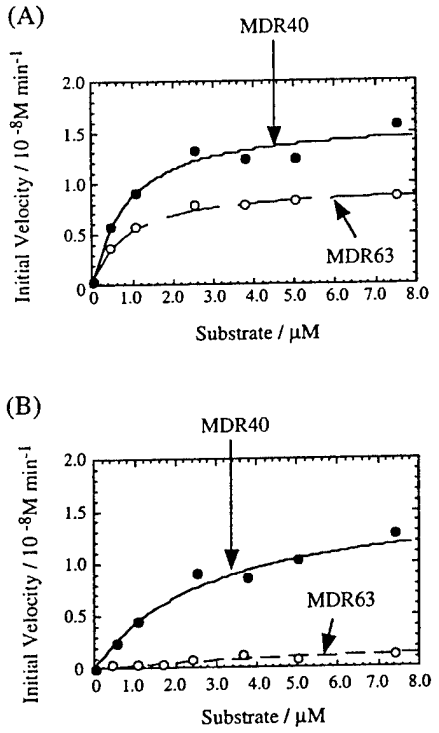


Figure 4. Examples of plots of initial rate of ribozyme cleavage at 37°C versus substrate concentration, under multiple turnover conditions. (**A**) Plots for the cleavage of wild-type substrates by MDR63 (open circle) and MDR40 (closed circle). (**B**) Plots for the cleavage of mis1 mismatched substrate by MDR 63 (open circle) and MDR 40 (closed circle). Curves were obtained with non-linear least square fits to Michaelis–Menten kinetics.

We also measured the kinetics under single turnover conditions. There are two protocols commonly used to initiate cleavage under single turnover conditions: one is initiation by combining ribozyme–substrate complex and Mg^{2+}. Another is initiation by combining the ribozyme–Mg^{2+} and the substrate–Mg^{2+} complexes. When the ribozyme–Mg^{2+} complex or the substrate–Mg^{2+} complex has unfavorable secondary structure, both saturated cleavage rate constants (k_{obs}) are the same (28). However, if there is a trap step or inhibition step that depends on secondary structure in the reaction mechanism, both saturated cleavage rate constants would not be the same. The stable ribozyme–Mg^{2+} complex or substrate–Mg^{2+} complex causes the saturated k_{obs} to decrease because of a small fraction of active complex (28,29). When we carried out the two experiments, the saturated k_{obs} (19mer substrate) by combining ribozyme–substrate complex with Mg^{2+} was 1.45 min^{-1} (data not shown). On the other hand, the saturated k_{obs} by combining the ribozyme–Mg^{2+} complex and the substrate–Mg^{2+} complex was 0.17 min^{-1}, suggesting that there is a trap step or inhibition step based on competition against binding of target in the reaction mechanism. The 0.17 min^{-1} value is in agreement with the k_{cat} value, 0.19 min^{-1}, measured for MDR63 under multiple turnover conditions. The k_{cat} value is also smaller than the predicted rate of product dissociation of 0.54 and 2.76 min^{-1} for both the products (30). These data show that the k_{cat} value for MDR63 under multiple turnover conditions is unlikely to be the rate of product dissociation because one only has to consider the steps up to the chemical step in single turnover conditions. Thus, the kinetic parameters with MDR63 under multiple turnover conditions contain the competition against binding of targets. Further, the catalytic efficiency of the minimal MDR40 hammerhead was found to be 3.6×10^5 M^{-1} min^{-1} under multiple turnover conditions. Thus, with the 19mer substrate RNA both the short ribozyme and the self-complementary, self-processed one have very similar activity under multiple turnover conditions. Therefore, we conclude that any self-binding that might be occurring with MDR63 does not adversely affect its *trans*-cleaving ability.

Specificity of MDR63 against mismatched substrates

To investigate sequence specificity for RNA cleavage by the MDR63 ribozyme, we prepared four mismatched 19 nt target RNAs (mis1, mis2, mis3 and mis4; Fig. 2). The mismatches were located within or near self-binding domains, two on each side of the catalytic domain. They were placed away from the cleavage site, however, because substitution near this site affects not only the substrate binding affinity but also the catalytic rate of cleavage (31). Again for comparison we studied the shortened MDR40 ribozyme.

Kinetics studies were then carried out under multiple turnover conditions with these two ribozymes and the complementary and mismatched 19mer target RNAs. Plots of the initial

Table 1. Kinetic parameters for RNA cleavage by the ribozymes[a]

Ribozyme	Substrate	k_{cat} (min^{-1})	K_m (μM)	$k_{cat}/K_m \times 10^{-5}$ (M^{-1} min^{-1})	Specificity[b]	MDR63/MDR40c	UG63/MDR40d
MDR40	Wild-type	0.32	0.88	3.6	1		
MDR40	mis1	0.32	2.6	1.2	0.33		
MDR40	mis2	0.33	2.3	1.4	0.39		
MDR40	mis3	0.91	4.8	1.9	0.53		
MDR40	mis4	0.76	4.9	1.5	0.42		
MDR63	Wild-type	0.19	0.74	2.6	1	0.74	
MDR63	mis1	0.05	8.2	0.06	0.02	0.05	
MDR63	mis2	0.04	8.2	0.05	0.02	0.03	
MDR63	mis3	0.54	5.7	0.95	0.36	0.50	
MDR63	mis4	0.20	3.7	0.54	0.21	0.36	
UG63	Wild-type	0.54	1.5	3.6	1		1
UG63	mis1	0.06	23.7	0.02	0.006		0.02
UG63	mis2	0.05	14.0	0.03	0.009		0.02
UG63	mis3	0.31	11.9	0.27	0.075		0.14
UG63	mis4	0.07	11.0	0.06	0.017		0.04

[a]All experiments were done in 50 mM Tris–HCl (pH 8.1) with 15 mM $MgCl_2$, 20 mM NaCl and 0.4 mM spermine at 37°C.
[b]Specificity = (k_{cat}/K_m with mismatched substrate)/(k_{cat}/K_m with wild-type substrate).
[c]MDR63/MDR40 = (k_{cat}/K_m with MDR63)/(k_{cat}/K_m with MDR40).
[d]UG63/MDR40 = (k_{cat}/K_m with UG63)/(k_{cat}/K_m with MDR40).

rates as a function of concentration showed hyperbolic shapes in all cases, suggesting that all ribozyme–substrate combinations behaved with Michaelis–Menten kinetics (Fig. 4). The kinetics data are given in Table 1. Specificity was defined as the ratio k_{cat}/K_m for mismatched and complementary targets. For the standard MDR40 ribozyme, specificities were 0.33, 0.39, 0.53 and 0.42 for mis1–mis4, respectively. Thus, these mismatched RNAs are cleaved with efficiencies one-third to one-half that of the complementary RNA by this standard ribozyme. Interestingly, the MDR63 ribozyme gave quite different results. Specificities were 0.02, 0.02, 0.36 and 0.21 for the same mismatches. Thus, the larger self-complementary ribozyme displays mismatch discrimination that is up to 20-fold greater than that of the ribozyme lacking this self-binding potential.

A single base change yields higher specificity

Examination of the specificity data with the MDR63 ribozyme shows that the two mismatches on one side of the ribozyme are much better discriminated than those on the other side (Table 1). Interestingly, the higher specificity occurred on the side predicted to have the stronger self-complementary hairpin sequence with 5 bp of possible duplex. We postulated that having stronger self-complementarity on the other side might possibly increase the sequence specificity in that *trans*-binding domain. To test this we designed a new self-processed ribozyme (UG63, Fig. 1B) having 1 bp of additional self-complementarity on the 5′-side of the ribozyme. This requires only a single U→G base change in the putative self-binding domain. Importantly, this mutation is located at a site where normal binding of substrate does not occur. Thus, if it affected cleavage it would be indirectly rather than by direct interaction with the target. RNA folding algorithms predict that the UG63 ribozyme can form a stronger hairpin, with a free energy of −3.5 kcal/mol, 3.1 kcal/mol more stable than the corresponding hairpin in MDR63. If self-binding is important in competing against mismatched target binding, then one would predict that higher sequence specificity against mis3 and mis4 might result.

As was done previously, a circular DNA encoding this second ribozyme was constructed. It was successfully transcribed *in vitro* by *E.coli* RNAP and gave products that appeared the same as those with the MDR63 circular vector (Fig. 3A). The fastest traveling band was isolated and was confirmed to be the UG63 mutant by RNase T1 sequencing (Fig. 3C).

Kinetics studies were carried out under conditions of multiple turnover for this second ribozyme using the same complementary and mismatched 19mer RNAs studied previously. The data are given in Table 1. The results show that the specificity of the UG63 ribozyme is considerably higher with the mis3 and mis4 mismatched RNAs, as compared to the previous MDR63 ribozyme. It cleaves them only 8 and 1.7% as efficiently as the complementary RNA, giving a 7- and 25-fold advantage over the standard 40mer ribozyme at these positions. Interestingly, it appears that specificity may also be slightly higher at the mis1 and mis2 mismatch positions; here the UG63 self-processed ribozyme has a 50- to 100-fold advantage over the short ribozyme. The high specificities for mis1 and mis2 are seen to depend mainly on K_m. Because of the extra base pair, the UG63 self-processed ribozyme has a greater energy penalty than MDR63 in binding to the substrate. Consistent with this, the K_m value for wild-type substrate with UG63 is larger than that for MDR63. The higher specificities

for mis1 and mis2 appear to be due to this energy penalty. The results suggest that competition between self-binding and target binding plays an important role in the sequence specificity of these self-processed ribozymes.

DISCUSSION

Previous studies of hammerhead ribozyme discrimination against mismatched targets have revealed levels of specificity similar to those of the present control MDR40 ribozyme. For example, when substrate RNAs were designed to have 8+7, 8+6 and 8+5 nt of complementarity for a hammerhead ribozyme, the specificities were 0.66, 0.66 and 0.5 for singly mismatched substrates, respectively (11). Mismatch discrimination is expected to be low when chemical cleavage is rate limiting, since most mismatches do not affect the active site geometry. If, on the other hand, product release is rate limiting, mismatched targets can be cleaved even more rapidly than complementary ones under multiple turnover conditions. Thus, minimal hammerhead ribozymes generally display quite low levels of sequence discrimination and the present self-binding strategy offers a substantial degree of improvement, with specificities of 0.075–0.006 for singly mismatched substrates with the UG63 ribozyme. This specificity appears to be due both to increases in K_m with a mismatch (by factors of 7- to 16-fold) and to decreases in k_{cat} with a mismatch (by 2- to 9-fold). It appears that both effects can be explained by self-binding of the RNA ends in UG63 relative to the MDR40 control (see discussion below).

Interestingly, this apparent self-binding does not inhibit the rate of substrate cleavage under multiple turnover conditions. Our data show that MDR63 and UG63 cleave a 19 nt wild-type substrate with a k_{cat}/K_m almost the same as the control MDR40. RNA folding algorithms predict a free energy (37°C) of –11.7 kcal/mol for the MDR63 folded structure with hairpins as in Figure 1 and –12.7 kcal/mol for UG63. On the other hand, the free energy for the complex between MDR63 or UG63 and the 19 nt target RNA is calculated to be –23.4 kcal/mol, using a catalytic core value of +3.3 kcal/mol (32). Thus, from a thermodynamic standpoint *trans*-cleavage is reasonable, since *trans*-binding is much more favorable than self-binding. From a kinetic standpoint, the k_{cat} values suggest that at 37°C the self-binding hairpins are melted at least as rapidly as the slowest step in the turnover cleavage mechanism. If one assumes a rate of 3.4×10^4 s^{-1} for self-annealing of one hairpin in MDR63 (33), then the rate constant for opening of this hairpin is expected to be 136 s^{-1}. This suggests that there are multiple chances for the target RNA to bind before the productive complex is formed. With a considerably stronger hairpin, however, one expects that the rate of cleavage would be slowed. Calculations suggest that binding domain hairpins longer than 9 bp in length (in this sequence context) might be kinetically inhibitory.

It is worth noting that the activity of MDR40 is low (by a factor of ~30) relative to a 'well-behaved' hammerhead ribozyme (HH16) studied by Hertel *et al.* (28). It has been observed previously that lower activities than predicted are seen in many ribozymes and they cannot always be simply explained (30). For example, when substrate RNAs were designed to have 8+8, 8+5 and 8+3 nt of complementarity for a previously studied hammerhead ribozyme, the k_{cat}/K_m values were 1.2×10^7, 0.5×10^7 and 1.4×10^7 M^{-1} min^{-1}, respectively (11). Further, when two hammerhead ribozymes were designed to have the same binding energies ($\Delta G°_{37} = -17.7$ and –17.6 kcal/mol) against the substrate RNAs, the k_{cat}/K_m values were 8.0×10^6 and 3.6×10^5 M^{-1} min^{-1}, respectively (11,21). This difference in k_{cat}/K_m values is similar in magnitude to that observed for MDR40 here. Such unpredicted differences depend on kinetically or thermodynamically inactive conformations in the reaction mechanism (11). For example, if the substrate binds in an alternative unreactive complex with the ribozyme, known as non-productive binding (34), the observed k_{cat} values decrease so that k_{cat}/K_m values also decrease. Since our 19 nt substrate has a UUUC sequence complementary to GAAA within the ribozyme core, it is possible that a non-productive complex lowers the overall activity for this sequence. Although the activity with MDR40 appears to be low, MDR63 and UG63 also have the same UUUC sequence. Therefore, the conditions with MDR40, MDR63 and UG63 are equal and MDR40 is expected to behave as a valid control.

MDR63 and UG63 display much greater abilities to distinguish mismatches than the standard MDR40 ribozyme and other hammerhead ribozymes in general (11). We observe that these high specificities depend both on increases in K_m and decreases in k_{cat} for mismatched targets. It is unclear whether the K_m is the same as the binding constant (K_d) of the ribozyme to the substrate because the K_d value was not determined by a separate method. From the standpoint of thermodynamic RNA folding predictions, however, the K_m effects can be understood by analysis of free energies of self-binding versus target binding. In the case of MDR40, the free energies for mis1 and mis2 target binding are calculated to be –17.5 and –17.8 kcal/mol (as compared to –21.7 for the complementary target). In the case of MDR63, the calculated free energy (–3.4 kcal/mol) for hairpin loop structures at self-complementary ends must be subtracted from these values and so the net free energy of binding is predicted to be –14.6 and –14.9 kcal/mol for mis1 and mis2, respectively (as compared to –18.8 kcal/mol for the complementary target). In fact, both experimentally determined K_m values of 8.2 μM for mis1 and mis2 with MDR63 are ~3.5-fold larger than those with MDR40. Similar effects are seen at the 5′-side, with the mis3 and mis4 targets. The free energy of target binding by MDR40 for these two is predicted to be –18.0 and –15.0 kcal/mol. The calculated net free energy for MDR63 is –15.1 and –12.1 kcal/mol. For UG63 it is –11.8 and –8.8 kcal/mol. As predicted by the free energy of target binding, the experimentally observed K_m value of 11.9 μM for mis3 with UG63 is the largest of the three ribozymes. The K_m value of 11.0 μM for mis4 with MDR63 is also the largest. Thus, the difference in K_m observed for these three ribozymes correlates with the calculated relative binding free energies with the RNA targets, suggesting that competition between self-binding and mismatched targets contributes to this sequence discrimination.

The observed high specificity of the self-processed ribozymes is not due to K_m effects alone. Unlike the short MDR40 ribozyme and most hammerhead ribozymes in general (11), the MDR63 and UG63 ribozymes display significantly decreased k_{cat} values with mismatched target RNAs. For the superior UG63 ribozyme, k_{cat} drops by as much as 180-fold with a mismatched target, while the MDR40 ribozyme shows either no effect or an increase in k_{cat}. If k_{cat} reflects only the

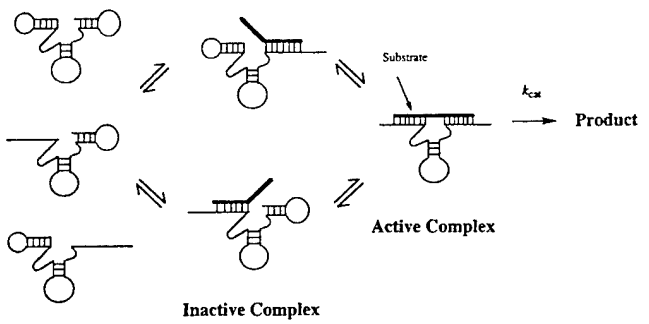


Figure 5. Schematic diagram of the fraying model, illustrating how self-complementarity in the ribozyme can increase specificity by competition between self-binding and binding of mismatched substrates. Self-binding is expected to increase the K_m with mismatched substrates by shifting the initial equilibrium away from the active complex. This is also expected to affect the observed rate of the chemical cleavage step (k_{cat}·[active complex]), again by lowering the concentration of the active complex, especially in the case of mismatched substrates.

chemical cleavage step, the decrease in k_{cat} might be due to alteration of the geometry near the cleavage site, although in the present case the positions of the mismatches were designed to have no effect on the catalytic rate of cleavage (31). In some cases the k_{cat} value reflects a rate limiting product release step, however, that seems an unlikely explanation in the present case, since the products are very similar for both short and self-processed ribozymes. Thus it appears that the results can be best explained by the fraying model for hammerhead ribozyme catalysis (35). In this model (Fig. 5) there are two non-productive ribozyme–target complexes that can be formed when only one side of the ribozyme binds the target (termed 'open states'). These non-productive complexes compete with productive complex formation (the 'closed state') and the observed k_{cat} value depends on the internal equilibrium constant for these complexes multiplied by the intrinsic k_{cat} value. Although the intrinsic k_{cat} is not affected by mismatches, the internal equilibrium between open and closed states certainly would be. In the case of the self-processed ribozymes, the internal equilibrium with mismatched targets would lie much more toward the side of the open complexes because of the competition with self-binding domains. Thus under the reaction conditions there is a much smaller fraction of active complex formed with mismatched targets than with complementary ones and this is reflected in the observed k_{cat} values.

Herschlag has carried out an analysis of specificity in ribozyme systems, leading to the prediction of strategies for increasing specificity between correct and incorrect substrates (36). One of these was the choice of A+U-rich targets, which permits substrate dissociation to be faster than chemical cleavage. A second was the use of high ionic strengths, which would increase the rate of equilibration between substrate and ribozyme. The third suggestion was to make use of unfavorable interactions between substrate and ribozyme, although a specific mechanism for carrying this out was not suggested. The present self-processed ribozymes also appear to utilize this third mechanism for increasing sequence specificity, decreasing substrate binding affinity by competition with self-structure. This strategy for hammerhead ribozymes has apparently not been tested previously. Interestingly, it was reported recently that the energetic penalty for conformational rearrangement with a group II intron ribozyme led to a reduction in the substrate binding energy so that the sequence specificity was enhanced (37,38). A related observation was made by Uhlenbeck, who observed that competition by secondary structure in a target RNA led to significant increases in sequence specificity of cleavage (33). The present study demonstrates how this concept can be successfully applied to ribozyme design, as opposed to target selection.

Although it was not carried out in a catalytic ribozyme system, another related observation merits mention in this context. Roberts and Crothers studied the specificity of binding of duplex DNA by a triplex-forming oligonucleotide and found that specificity could be increased by designing a self-binding domain (termed a 'stringency clamp') into the oligonucleotide (39). This was important in demonstrating the value of competition between self-binding and target binding in a hybridization experiment.

ACKNOWLEDGEMENTS

We thank the DOD Breast Cancer Research Program, administered by the US Army and the National Institutes of Health (GM46625) for support of this work.

REFERENCES

1. Buzayan,J.M., Gerlach,W.L. and Bruening,G. (1986) *Nature*, **323**, 349–353.
2. Hutchins,C.J., Rathjen,P.D., Forster,A.C. and Symons,R.H. (1986) *Nucleic Acids Res.*, **14**, 3627–3640.
3. Prody,G.A., Bakos,J.T., Buzayan,J.M., Schneider,I.R. and Bruening,G. (1986) *Science*, **231**, 1577–1580.
4. Forster,A.C. and Symons,R.H (1987) *Cell*, **49**, 211–220.
5. Uhlenbeck,O.C. (1987) *Nature*, **328**, 596–600.
6. Haseloff,J. and Gerlach,W.L. (1988) *Nature*, **334**, 585–591.
7. Zhou,D.-M. and Taira,K. (1998) *Chem. Rev.*, **98**, 991–1026.
8. Eckstein,F. and Lilley,D.M.J. (eds) (1996) *RNA Catalysis: Nucleic Acids and Molecular Biology*. Springer Verlag, Berlin, Germany, Vol. 10.
9. Chen,H., Ferbeyre,G. and Cedergren,R. (1997) *Nature Biotechnol.*, **15**, 273–277.
10. Kuwabara,T., Warashina,M., Nakayama,A., Ohkawa,J. and Taira,K. (1999) *Proc. Natl Acad. Sci. USA*, **96**, 1886–1891.
11. Hertel,K.J., Herschlag,D. and Uhlenbeck,O.C. (1996) *EMBO J.*, **15**, 3751–3757.
12. Daubendiek,S.L., Ryan,K. and Kool,E.T. (1995) *J. Am. Chem. Soc.*, **117**, 7818–7819.
13. Daubendiek,S.L. and Kool,E.T. (1997) *Nature Biotechnol.*, **15**, 273–277.
14. Diegelman,A.M. and Kool,E.T. (1998) *Nucleic Acids Res.*, **26**, 3235–3241.
15. Diegelman,A.M. and Kool,E.T. (1999) *Chem. Biol.*, **9**, 569–576.
16. Usman,N., Ogilvie,K.K., Jiang,M.-V. and Cedergren,R. (1987) *J. Am. Chem. Soc.*, **109**, 7845–7854.
17. Diegelman,A.M. and Kool,E.T. (1999) In Glick,G.D. (ed.), *Current Protocols in Nucleic Acid Chemistry*. John Wiley & Sons, New York, NY, in press.
18. Richards,E.G. (1975) In Fasman,G.D. (ed.), *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*, 3rd Edn. CRC Press, Cleveland, OH, Vol. I, p. 197.
19. Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.C. (1987) *Nucleic Acids Res.*, **15**, 8783–8798.
20. Hendry,P., McCall,M.J. and Lockett,T.J. (1997) In Turner,P.C. (ed.), *Methods in Molecular Biology*. Humana Press, Totowa, NJ, Vol. 74.
21. Fedor,M.J. and Uhlenbeck,O.C. (1992) *Biochemistry*, **31**, 12042–12054.
22. Kool,E.T. (1998) *Acc. Chem. Res.*, **31**, 502–510.
23. Kobayashi,H., Dorai,T., Holland,J.F. and Ohnuma,T. (1994) *Cancer Res.*, **54**, 1271–1275.

24. Turner,D.H., Sugimoto,N. and Freier,S.M. (1988) *Annu. Rev. Biophys. Biophys. Chem.*, **17**, 167–192.
25. Serra,M.J. and Turner,D.H. (1995) *Methods Enzymol.*, **259**, 242–261.
26. Xia,T., SantaLucia,J.,Jr, Burkard,M.E., Kierzek,R., Schroeder,S.J., Jiao,X., Cox,C. and Turner,D.H. (1998) *Biochemistry*, **37**, 14719–14735.
27. Mathews,D.H., Sabina,J., Zuker,M. and Turner,D.H. (1999) *J. Mol. Biol.*, **288**, 911–940.
28. Hertel,K.J., Herschlag,D. and Uhlenbeck,O.C. (1994) *Biochemistry*, **33**, 3374–3385.
29. Walstrum,S.A. and Uhlenbeck,O.C. (1990) *Biochemistry*, **29**, 10573–10576.
30. Stage-Zimmermann,K.T. and Uhlenbeck,O.C. (1998) *RNA*, **4**, 875–889.
31. Werner,M. and Uhlenbeck,O.C. (1995) *Nucleic Acids Res.*, **23**, 2092–2096.
32. Hertel,K.J., Stage-Zimmermann,T.K., Ammons,G. and Uhlenbeck,O.C. (1998) *Biochemistry*, **37**, 16983–16988.
33. Pörschke,D. (1974) *Biophys. Chem.*, **1**, 381–384.
34. Fersht,A. (1985) *Enzyme Structure and Mechanism*, 2nd Edn. W.H. Freeman and Co., New York, NY.
35. Hertel,K.J., Peracchi,A., Uhlenbeck,O.C. and Herschlag,D. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 8497–8502.
36. Herschlag,D. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 6921–6925.
37. Qin,P.Z. and Pyle,A.M. (1999) *J. Mol. Biol.*, **291**, 15–27.
38. Xiang,Q., Qin,P.Z., Michels,W.J., Freeland,K. and Pyle,A.M. (1998) *Biochemistry*, **37**, 3839–3849.
39. Roberts,R.W. and Crothers,D.M. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 9397–9401.

Current Opinion in Chemical Biology **2000, in press.**

Synthetically modified DNAs as substrates for polymerases

Eric T Kool

DNA polymerase enzymes process their natural substrate with very high specificity. Yet recent experiments have shown that these enzymes can also process DNA in which the backbone or bases are modified to a surprising degree. Such experiments have important implications in understanding the mechanisms of DNA replication, and suggest important biotechnological uses as well.

Address
Department of Chemistry, Stanford University, Stanford, CA 94305, USA
e-mail: kool@stanford.edu

Abbreviations

dNTP deoxyribonucleoside 5'-triphosphate

rNTP ribonucleoside 5'-triphosphate

Kf Klenow fragment of *E. coli* DNA Polymerase I

Introduction

DNA polymerase enzymes are inextricably linked to biology, since the very definition of biological organisms requires that they copy themselves. They are therefore probably one of the earlier classes of protein enzymes to evolve, and have had a long time to become efficient at what they do. Several groups of chemical researchers have recently taken up the challenge of finding new molecular designs that can trick polymerases into accepting and replicating them. This is a difficult problem to tackle: these enzymes are highly specific, making less than one error in a million nucleotide additions. They make extensive contact with the DNA acting as a template, holding it over 4-5 base pairs, and form a sterically tight active site pocket with several hydrogen-bonded contacts.

Why find new substrates for DNA polymerases? There are several answers to this question, which is why an increasing number of research groups have entered the field. First, functioning analogs of DNA can provide useful insights into the basic mechanisms by which polymerases work. Second, DNA replication is central to biotechnology, and finding new substrates can have important applications in amplifying genetic sequences, in genetic therapies, in molecular diagnostics, in combinatorial *in vitro* selection strategies, and in further-reaching biological applications such as expansion of the genetic code.

The chemistry of a DNA polymerase involves phosphodiester bond formation (Figure 1). The nucleophile in the reaction is the 3' hydroxyl group at the end of the growing (primer) strand of DNA. This primer strand is bound in a double helix to a longer template strand, which encodes the chemical information that is being transferred to the strand being synthesized. The leaving group in the nucleophilic substitution reaction is pyrophosphate, which is originally attached to an incoming nucleotide as part of a triphosphate group. The site of attack is phosphorus of the alpha-phosphate in the triphosphate (the one closest to the sugar), which undergoes inversion of stereochemistry. For correctly matched natural DNA nucleotides, the rate-limiting step (for those

polymerases studied in detail) is not this bond formation, but rather a conformational change preceding this chemical step. For incorrectly matched nucleotides, the chemistry step slows to the point where it becomes rate limiting. It should be noted that whether a modified substrate is accepted by a polymerase is not a binary, yes-or-no proposition. Measurable polymerase activities can vary over more than seven orders of magnitude of k_{cat}/K_M values. Thus, quantitative data can be very useful in making judgements of how well a modification is accepted. It might also be noted that polymerases often have other enzymatic activities, including hydrolytic removal of terminal nucleotides in 3' exonuclease and 5' exonuclease domains, but these will not be discussed here.

Before proceeding further it should be stressed that there are many different DNA polymerases. While a virus may need only one such enzyme to be replicated, higher organisms utilize a greater number of more specialized enzymes. For example, humans are known to have several DNA polymerases, each with a very specific role to play in DNA replication and repair. Some are highly processive and proceed with high fidelity, while others make DNA slowly and are much more error-prone. Thus, not all polymerases are exactly alike, and so conclusions of chemical studies with one enzyme may not necessarily be applied to all of them. With that caution, the following is a brief review of recent work in this growing field. Note that some of the earlier work has been summarized in reviews recently [1•-3], so here we focus mainly on findings presented in the last one to two years.

Backbone Modifications

Chemical alterations to phosphodiesters

To some extent, DNA polymerases can accept chemical modifications either to the template strand or to the new, growing strand. In the former case the modifications are put in place by the researcher using chemical synthesis, and one asks the polymerase to read and replicate this modified structure. In the latter, synthetically modified nucleoside triphosphates (dNTPs for DNA or rNTPs for RNA) are put into the new strand by the

enzyme. If one desires newly synthesized DNA that is chemically altered in some way, then there are three sites of the dNTP where modifications can be made: in the alpha-phosphate group, in the sugar, and in the nucleobase. The first two are backbone modifications, and few reports exist on sugar modifications that function well with DNA polymerases. One example is a report of acyclic sugar replacements, which are reportedly accepted by the Kf enzyme as substrate dNTPs [4]. They are inserted at the end of the growing strand; however, they act as terminators, which suggests that an interaction further into the active site is missing or unfavorable. In a related vein are studies with nucleotides having substituents at the 4' position of the sugar [5-7•]. These nucleotides are in some cases active substrates for polymerases, but they tend to act as chain terminators, depending on the polymerase. It should be noted, however, that DNA synthesis terminators can be extremely useful; for example, dideoxynucleotide terminators are used in standard DNA sequencing methods. The terminators of Giese have been useful as probes of steric interactions in the active site of HIV-1 reverse transcriptase [7•].

There are at least three very interesting synthetic alterations of the alpha-phosphate of dNTPs that are readily accepted by DNA polymerases. First is the thio replacement of one of the non-bridging oxygens (only the Sp oxygen can be replaced) [8]. This replacement is very efficiently accepted by Taq or Kf enzymes, and produces stereoregular (Rp) phosphorothioate DNA as a product [8,9]. Second is a very interesting report that this non-bridging oxygen can be replaced by a methyl group [10]. The product of this synthesis of non-charged, stereoregular methylphosphonate DNA; it is remarkable that this lack of charge can be accepted so readily. Quantitative kinetic data are as yet unavailable for this modification. Finally, the surprising replacement of BH_3- for O- can be made at the alpha triphosphate, giving boronate DNA as a product [11•,12]. Using the Taq enzyme, this occurs with efficiency essentially the same as natural nucleotides, which is remarkable.

There are few studies of backbone modifications in the template strand. One such example is a replacement of the 5' bridging oxygen of the phosphodiester. One of these groups appears to be qualitatively very well tolerated by the Klenow enzyme, with complete replication up to and past this site occurring without pauses [13]. Quantitative information is not yet available for this modification.

Structural alterations: circular DNAs

A different type of modification of the template backbone is in its overall shape. Different topological forms, such as circles, catenanes, or knots, might well give quite different results with polymerases. While to our knowledge synthetically made DNA knots have not yet been replicated, small circular single-stranded DNAs have come under close scrutiny recently as substrates for polymerase enzymes [14-16]. One of the reasons for this interest is that a polymerase reading a circular template makes repeated copies of the encoded sequence without dissociation or thermal heat-cool cycles. Thus it is a potentially useful amplification system. This "rolling circle replication" (abbreviated RCA and RCR) has been used to detect single copies of genetic sequences [17]. What is especially surprising about this is that quite small circles can be effective substrates. For example, it was shown that 34mer DNA circles are very efficient templates [16], and a recent paper established that even a cyclic 13mer acts to some degree as a substrate for common polymerases such as Kf [18].

Also noteworthy is the finding that such small synthetic DNA circles can act as efficient templates for RNA polymerases as well [15]. This "rolling circle transcription" (RCT) can be used to generate self-processed catalytic RNAs in quite pure form and with high levels of amplification [19,20]. This can be made to occur inside bacterial cells (Ohmichi T, Kool ET, unpublished), and it is conceivable that such nanometer-sized circles might one day encode biologically active RNAs inside human cells as well. Once again, even a 13mer circle works, to a small extent, as a template for T7 RNA polymerase [18].

Base modifications

Altered Watson-Crick hydrogen bonding arrangements

One ongoing area of research has been aimed at finding ways to get DNA polymerases to act nonspecifically. Several groups have worked on development of hydrogen-bonding-altered nucleotides that, when present in a template, will nonspecifically direct incorporate any one of the four natural bases into the growing primer strand [21]. Nonspecifically-acting nucleoside triphosphates have also been described [22]. This would have useful applications in mutagenesis and other applications. To date, such "universal base" analogs have been successful in directing mixtures but have not yet reached the stage of directing equal amounts of all four bases.

Early research on DNA polymerases operated under the assumption that complementarity of Watson-Crick hydrogen bonds was an essential part of maintaining replication activity. It is now known that having such bonds is not an absolute requirement; however, a quite viable approach to altering DNA bases involves altering the arrangement of donors and acceptors. As pointed out by Benner, there are eight possible arrangements of donors and acceptors in a triply-H-bonded pair, and only four of these arrangements are used in the natural bases [23]. The Benner laboratory has synthesized and studied a number of the nonnatural cases. One of the most successful is the isoG-isoC pair (Figure 2), which was first noted by Rich nearly four decades ago [24]. This pair is formed by DNA polymerases with qualitatively high efficiency, although it appears that extension is hindered somewhat. It is possible that this hesitation arises from unsatisfied hydrogen bonds between the polymerase and the DNA minor groove [25,26]. The pair works with a good degree of specificity even in the presence of the other four bases; although there is some interference since isoG miscodes for T quite frequently. The 5-methyl substitution of isoC ameliorates an earlier problem of hydrolytic instability seen for unsubstituted isoC [27]. The isoC base can serve as a template for RNA polymerases as well, directing insertion of ribo(isoG) triphosphate [28]. Finally, the Chamberlin and Benner

laboratories showed in a remarkable experiment that this pair can be used to expand the genetic code by functioning in the codon-anticodon interaction, directing incorporation of a nonnatural amino acid into a protein in vitro [29].

Other potentially replicable pairs, such as one termed kappa-pi, have also been studied by the Benner laboratory [30,31] (Figure 2). It also functions to some extent with DNA polymerases, although it has not seen the number of applications that the isoC-isoG pair has. Benner has performed a survey of a number of different polymerase enzymes with such modified nucleotides [30,31]; from such results it has become clear that different enzymes vary quite significantly in their ability to handle nonnatural bases.

Recent studies have utilized a combined strategy involving steric effects and hydrogen bonding groups (Figure 2). Pyridinones are have been utilized in templates and as nucleoside triphosphates with signficant success [32,33]. Some early work on a closely related approach also merits mentioned in this context. Rappaport described a modified base pair that functions quantitatively quite well. This pair, 5-methyl-2-pyrimidinone / 6-thioguanine (Gs), is an altered version of C-G. The pair is processed with significant selectivity, although it suffers from interfering misinsertion of unmodified C opposite Gs [34•].

Conjugates of bases

The most widespread use of modified bases with polymerases occurs in DNA sequencing and labeling [35,36]. When fluorescent-labeled nucleotides are used, the base must be modified with a linker to a fluorophore. It is now clear that the structure (geometry and length) of the linker and conjugate can have strong effects on polymerase activity [36,37]. The most common substitution structure involves alkynyl [38] or alkenyl chains directly attached to a carbon of the DNA base. Moreover, it is also clear that the position of substitution of the four natural DNA bases has a strong effect as well. In general, for pyrimidines (C and T) the 5-position is preferred by far. The closest analogous position

on purines is the 7-position; for substitution here, 7-deaza analogs are used. Even after this optimization, however, such analogues are often relatively poor substrates, sometimes orders of magnitude less efficiently processed than natural nucleotides. There is not sufficient space here to go into further detail on the earlier work on such conjugates, which in any case has mostly been published prior to last year.

One related research area of high interest has been the use of nucleotides that maintain Watson-Crick hydrogen bonding but which are altered in other ways, such as by deletion or addition of single H-bonding groups [39-42]. Base analogs in this class include 2-aminopurine, 2,6-diaminopurine, 7-deazaadenine, and several others. These analogs often can be useful in probing noncovalent contacts between DNAs, RNAs, and/or proteins. One particularly impressive recent development in this area has been the nucleic acid interference mapping (NAIM) methodology of Strobel [42•].

Nonpolar, non-H-bonding bases

Although Watson-Crick hydrogen bonds do provide some stabilization to the DNA helix, it was recently established that they are not necessary for high efficiency replication of base pairs by DNA polymerases. This was first realized when it was discovered that difluorotoluene deoxynucleoside (dF) (Figure 2), a nonpolar molecule that mimics thymidine in structure and conformation [43], is replicated with high efficiency by DNA polymerases such as Kf and T7 [44,45]. It serves in a template very well, directing insertion of adenine with quantitative efficiency and selectivity approaching those of the natural thymidine. In addition, the reverse is true; the nucleoside triphosphate analog (dFTP) is inserted into a growing primer with high efficiency and selectivity [46]. This molecule can also be replicated quite efficiently opposite a nonpolar surrogate of adenine such as 4-methylbenzimidazole deozynucleoside (dZ) [47,48] or an aza-substituted variant thereof (dQ) [26]. Presumably, base stacking of these nonpolar molecules helps compensate for the lack of hydrogen bonds in helping the nucleotide to bind in the active site [49]. This work led to the proposal that steric exclusion alone in the tight active site

can serve as an important factor in selectivity of DNA replication [44,45,50]. In this model, the template base fills approximately half of the site, leaving the rest of the space unoccupied. This template base thus determines the shape of the space remaining to be filled, and the two bases together must fit into a shape resembling that of the DNA double helix. Note that variations in tightness of fit may alter specificity of different enzymes [51].

Subsequent to this it was shown that not only are hydrogen bonds not needed for replication, but neither are shapes resembling the natural pyrimidine and purine skeletons, as long as the shapes are complementary. For example, an abasic sugar ((dϕ), where the "base" is essentially a proton or a hydroxyl group) can act as an efficient template nucleotide when the enzyme is offered dPTP, a nucleoside triphosphate with the large aromatic compound pyrene replacing the DNA base [52•] (Figure 2). The pair is formed with high selectivity and with efficiency similar to that of a natural base pair, using either the Kf enzyme or T7 DNA polymerase. Models of DNA suggest that pyrene fits comfortably in a DNA helix opposite dϕ, and experiments have confirmed that this pair is actually pairs selectively and stably in DNA [53]. The findings suggest that, as long as two bases or analogs fit opposite one another in a standard helix, and stack with good affinity, a polymerase may process the pair (at least for the initial insertion step). The main exception to this is when one of the two is highly polar and the other, nonpolar, which incurs the energetic cost of desolvating the polar partner [50]. For example, nonpolar dZ is replicated well opposite dF, but efficiency is significantly lower for dZ opposite highly polar dT.

More recent results have confirmed that hydrogen bonding is not a requirement for high polymerase activity. Schultz and Romesberg have reported an impressive array of additional aromatic, nonpolar bases that make stable and selective pairs, and can be formed by polymerases as well [54•,55]. One of the most interesting examples is a nucleoside built from propynylisocarbostyril (Figure 2) that forms a self-pair. In

synthetic DNAs this compound strongly prefers to pair with itself rather than natural bases, and the pair appears to be more stabilizing even than a native G-C pair. When tested with polymerases, this compound is inserted opposite another copy of itself with high quantitative efficiency, and it can even function in the presence of all four of the other base pairs.

One current limitation of the pyrene-abasic pair and the isocarbostyril self-pair is that, although they are formed efficiently, they serve as stops to further replication (the extension step). While terminators do have some useful applications, many possible applications require nucleotides that allow complete replication of DNA strands. One of the most important factors in successful DNA synthesis after a given pair is initially made, is that hydrogen bonds between the polymerase and the DNA minor groove be satisfied [25,26]. This was clearly demonstrated recently using nonpolar nucleoside isosteres of thymine and adenine. It was shown that the Kf enzyme forms a kinetically important hydrogen bond at one specific site in the template strand. Nonpolar analogs that can form this hydrogen bond function very well in continuing DNA synthesis. However, the hydrogen bonding is not the only factor in this extension step. The Schultz and Romesberg groups have observed that the isocarbostyril self-pair, which does have a potential minor groove H-bond acceptor, still acts as a terminator [54]. Thus it is likely that the geometry and steric fit of the pair in the active site is also important to continuing synthesis after a pair is put in place.

Conclusions and Future Prospects

These recent findings have important implications in the mechanism of how polymerases function. First, Watson-Crick hydrogen bonding groups can be rearranged or even deleted. It is clear that hydrogen bonds are not necessary for high efficiency base pair synthesis by at least some of these enzymes. However, with native enzymes it does appear that at least one minor groove hydrogen bond between polymerase and DNA is needed for continuing DNA synthesis with high efficiency. In addition, not all non-H-

bonded bases have demonstrated the high selectivity that is observed for replication of normal bases. Thus, it appears that the Watson-Crick hydrogen bonding groups may add a significant degree of selectivity beyond what steric shape alone can afford. It seems likely that molecules such as the nonpolar nucleoside isosteres, which maintain natural DNA structure [56], will serve as generally useful probes of protein-DNA interactions. In addition, new nonnatural base pairs may offer a number of useful applications in genetic typing, in DNA sequencing, and in encoding new amino acids into proteins.

From the studies of circular modified DNAs it also appears that DNA polymerases can tolerate a surprising degree of structural distortion. The use of topologically closed molecules as templates offers a new way to engender amplification of genetic signals and biologically active sequences. It also appears that in some cases, and in some applications, circular DNAs may behave as better templates than the classical linear DNA templates that are commonly used by researchers.

Acknowledgements

I thank the National Institutes of Health (GM52956) and the Army Breast Cancer Research Program for their support of our work in the use of modified DNAs as substrates for polymerases. I also gratefully acknowledge the Biological Division of the American Chemical Society and Pfizer Inc. for supporting the 2000 Pfizer Award in Enzyme Chemistry. Finally, I thank my co-workers, whose research is cited here.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

• of special interest

•• of outstanding interest

1. • Kool ET, Morales JC, Guckian KM: **Mimicking the structures and functions of DNA: insights into DNA stability and replication.** *Angew Chem Int Ed* 2000, **39**:990-1009.

Recent detailed review of issue of noncovalent bonding in polymerase functions.

2. Benner SA, Battersby TR, Eschgfaller B, Hutter D, Kodra JT, Lutz S, Arslan T, Baschlin DK, Blattler M, Egli M, Hammer C, Held HA, Horlacher J, Huang Z, Hyrup B, Jenny TF, Jurczyk SC, Konig M, von Krosigk U, Lutz MJ, MacPherson LJ, Moroney SE, Muller E, Nambiar KP, Piccirilli JA, Switzer CY, Vogel JJ, Richert C, Roughton AL, Schmidt J, Schneider KC, Stackhouse J: **Redesigning nucleic acids.** *Pure Appl Chem* 1998, **70**:263-266.
3. Kool ET: **Recognition of DNA, RNA, and proteins by circular oligonucleotides.** *Accounts Chem Res* 1998, **31**:502-510.
4. Martinez CI, Ansari MA, Gibbs R, Burgess K: **Acyclic nucleoside triphosphate analogs as terminators in biocatalytic DNA replication.** *Bioorg Med Chem Lett* 1997, **7**:3013-3016.
5. Hess MT, Schwitter U, Petretta M, Giese B, Naegeli H: **DNA synthesis arrest at C4'-modified deoxyribose residues.** *Biochemistry* 1997, **36**:2332-2337.
6. Marx A, MacWilliams MP, Bickle TA, Schwitter U, Giese B: **4'-acylated thymidines: a new class of DNA chain terminators and photocleavable DNA building blocks.** *J Am Chem Soc* 1997, **119**:1131-1132.
7. • Marx A, Spichty M, Amacker M, Schwitter U, Hübscher U, Bickle TA, Maga G, Giese B: **Probing interactions between HIV-1 reverse transcriptase and its DNA substrate with backbone-modified nucleotides.** *Chem Biol* 1999, **6**:111-116.

Thorough analysis of backbone steric interactions with HIV-1 reverse transcriptase.

8. Brody RS, Frey PA: **Unambiguous determination of the stereochemistry of nucleotidyl transfer catalyzed by DNA polymerase I from Escherichia coli.** *Biochemistry* 1981, **20**:1245-1252.

9. Burgers PM, Eckstein F: **A study of the mechanism of DNA polymerase I from Escherichia coli with diastereomeric phosphorothioate analogs of deoxyadenosine triphosphate.** *J Biol Chem* 1979, **254**:6889-6893.

10. Dineva MA, Chakurov S, Bratovanova EK, Devedjiev I, Petkov DD: **Complete template-directed enzymatic synthesis of a potential antisense DNA containing 42 methylphosphonodiester bonds.** *Bioorg Med Chem* 1993, **1**:411-414.

11. Shaw BR, Sergueev D, He KZ, Porter K, Summers J, Sergueeva Z, Rait V: **Boranophosphate backbone: a mimic of phosphodiesters, phosphorothioates, and methyl phosphonates.** *Methods Enzymol* 2000, **313**:226-257.

Description of remarkable backbone alteration that functions with polmerase enzymes.

12. He KZ, Porter KW, Hasan A, Briley JD, Shaw BR: **Synthesis of 5-substituted 2 '-deoxycytidine 5 '-(alpha-P-borano)triphosphates, their incorporation into DNA and effects on exonuclease.** *Nucleic Acids Res* 1999, **27**:1788-1794.

13. Xu Y, Kool ET: **Chemical and enzymatic properties of 5'-bridging phosphorothioester linkages in DNA.** *Nucleic Acids Res* 1998, **26**:3159-3164.

14. Fire A, Xu SQ: **Rolling replication of short DNA circles.** *Proc Natl Acad Sci USA* 1995, **92**:4641-4645.

15. Daubendiek SL, Ryan K, Kool ET: **Rolling circle RNA synthesis: circular oligonucleotides as efficient substrates for T7 RNA polymerase.** *J Am Chem Soc* 1995 **117**:7818-7819.

16. Liu D, Daubendiek SL, Zillmann MA, Ryan K, Kool ET: **Rolling circle DNA synthesis: small circular oligonucleotides as efficient templates for DNA polymerases.** *J Am Chem Soc* 1996, **118**:1587-1594.

17. Lizardi PM, Huang X, Zhu Z, Bray-Ward P, Thomas DC, Ward DC: **Mutation detection and single-molecule counting using isothermal rolling-circle amplification.** *Nat Genet* 1998 **19**:225-232.

18. Frieden M, Pedroso E, Kool ET: **Tightening the belt on polymerases: Evaluating the physical constraints on enzyme substrate size.** *Angew Chem Int Ed* 1999, **38**:3654-3657.

19. Daubendiek S, Kool ET: **Generation of catalytic RNAs by rolling transcription of synthetic DNA nanocircles.** *Nature Biotechnol* 1997, **15**:273-277.

20. Diegelman AM, Kool ET: **Generation of circular RNAs and catalytic RNAs by transcription of circular DNA oligonucleotides encoding hairpin ribozymes.** *Nucleic Acids Res* 1998, **26**:3235-3241.

21. Hoops GC, Zhang PM, Johnson WT, Paul N, Bergstrom DE, Davisson VJ: **Template directed incorporation of nucleotide mixtures using azole-nucleobase analogs.** *Nucleic Acids Res* 1997, **25**:4866-4871.

22. Hill F, Loakes D, Smith CL, Williams DM, Brown DM: **Synthesis and polymerase incorporation properties of a tricyclic pyrrolopyrimidine related to N-6-hydroxy-2 '-deoxyadenosine.** *Nucleosides Nucleotides* 1999, **18**:573-574.

23. Piccirilli JA, Krauch T, Moroney SE, Benner SA: **Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet.** *Nature* 1990, **343**:33-37.

24. Rich, A: In **Horizons in Biochemistry**; Kasha, M. Pullman, B., Eds.; Academic Press: New York, 1962; pp 103-126.

25. Bebenek K, Beard WA, Darden TA, Li L, Prasad R, Luton BA, Gorenstein DG, Wilson SH, Kunkel TA: **A minor groove binding track in reverse transcriptase.** *Nature Struct Biol* 1997, **4**:194-197.

26. Morales JC, Kool ET: **Minor groove interactions between polymerase and DNA: More essential than Watson-Crick bonds?** *J Am Chem Soc* 1999, **121**:2723-2724.

27. Jurczyk SC, Kodra JT, Park JH, Benner SA, Battersby TR: **Synthesis of 2 '-deoxyisoguanosine 5 '-triphosphate and 2 '-deoxy-5-methylisocytidine 5 '-triphosphate.** *Helv Chim Acta* 1999, 82:1005-1015.

28. Tor Y, Dervan PB: **Site-specific enzymatic incorporation of an unnatural base, N(6)-(6-aminohexyl)isoguanosine, into RNA.** *J Am Chem Soc* 1993, **115**:4461-4467.

29. Bain JD, Switzer C, Chamberlin AR, Benner SA: **Ribosome-mediated incorporation of a non-standard amino acid into a peptide through expansion of the genetic code.** *Nature* 1992, **356**:537-539.

30. Horlacher J, Hottiger M, Podust VN, Hubscher U, Benner SA: **Recognition by viral and cellular DNA polymerases of nucleosides bearing bases with nonstandard hydrogen bonding patterns.** *Proc Natl Acad Sci USA* 1995, **92**:6329-6333.

31. Lutz MJ, Held HA, Hottiger M, Hubscher U, Benner SA: **Differential discrimination of DNA polymerase for variants of the non-standard nucleobase pair between xanthosine and 2,4-diaminopyrimidine, two components of an expanded genetic alphabet.** *Nucleic Acids Res* 1996, **24**:1308-1313.

32. Ishikawa M, Hirao I, Yokoyama S: **Synthesis of 3-(2-deoxy-β-D-ribofuranosyl)pyridin-2-one and 2-amino-6-(N,N-dimethylamino)-9-(2-deoxy- β -D-ribofuranosyl)purine derivatives for an unnatural base pair.** *Tetrahedron Lett* 2000, **41**:3931-3934.

33. Hirao I, Ohtsuki T, Mitsui T, Yokoyama S: **Dual specificity of the pyrimidine analogue, 4-methylpyridin-2-one, in DNA replication.** *J Am Chem Soc* 2000, **122**:6118-6119.

34. Rappaport HP: **Replication of the base pair 6-thioguanine/5-methyl-2-pyrimidine with the large Klenow fragment of Escherichia coli DNA polymerase I.** *Biochemistry* 1993, **32**:3047-3057.

35. Langer PR, Waldrop AA, Ward DC: **Enzymatic synthesis of biotin-labeled polynucleotides: novel nucleic acid affinity probes.** *Proc Natl Acad Sci USA* 1981, **78**:6633-6637.

36. Zhu Z, Chao J, Yu H, Waggoner AS: **Directly labeled DNA probes using fluorescent nucleotides with different length linkers.** *Nucleic Acids Res* 1994, **22**:3418-3422.

37. Brandis JW: **Dye structure affects Taq DNA polymerase terminator selectivity.** *Nucleic Acids Res* 1999, **27**:1912-1918.

38. Battersby TR, Ang DN, Burgstaller P, Jurczyk SC, Bowser MT, Buchanan DD, Kennedy RT, Benner SA: **Quantitative analysis of receptors for adenosine nucleotides obtained via in vitro selection from a library incorporating a cationic nucleotide analog.** *J Am Chem Soc* 1999, **121**:9781-9789.

39. Bailly C, Payet D, Travers AA, Waring MJ: **PCR-based development of DNA substrates containing modified bases: An efficient system for investigating the role of the exocyclic groups in chemical structural recognition by minor groove binding drugs and proteins.** *Proc Natl Acad Sci USA* 1996, **93**:13623-13628.

40. Guo MJ, Hildbrand S, Leumann CJ, McLaughlin LW, Waring MJ: **Inhibition of DNA polymerase reactions by pyrimidine nucleotide analogues lacking the 2-keto group.** *Nucleic Acids Res* 1998, **26**:1863-1869.

41. Hacia JG, Woski SA, Fidanza J, Edgemon K, Hunt N, McGall G, Fodor SP, Collins FS: **Enhanced high density oligonucleotide array-based sequence analysis using modified nucleoside triphosphates.** *Nucleic Acids Res* 1998, **26**:4975-4982.

42. Strobel SA: **A chemogenetic approach to RNA function/structure analysis.**

• *Curr Opin Struct Biol* 1999, **9**:346-352.

A useful new RNA mapping strategy is described.

43. Guckian KM, Kool ET: **Highly precise shape mimicry by difluorotoluene**

deoxynucleoside, a replication-competent substitute for thymidine. *Angew Chem Int Ed* 1998, **36**:2825-2828.

44. Moran S, Ren RXF, Rumney S, Kool ET: **Difluorotoluene, a nonpolar isostere of thymine, codes specifically and efficiently for adenine in DNA replication.** *J Am Chem Soc* 1997, **119**:2056-2057.

45. Moran S, Ren RXF, Kool ET: **A thymidine triphosphate shape mimic lacking Watson-Crick pairing ability is replicated with high specificity.** *Proc Natl Acad Sci USA* 1997, **94**:10506-10511.

46. Liu D, Moran S, Kool ET: **Bistranded, multisite replication of an unconventional difluorotoluene-adenine base pair and confirmation by "inverse" sequencing.** *Chemistry & Biology* 1997, **4**:919-926.

47. Morales JC, Guckian KM, Sheils CJ, Kool ET: **Structure and base pairing properties of a replication-competent nonpolar isostere for deoxyadenosine.** *J Org Chem* 1998, **63**:9652-9656.

48. Morales JC, Kool ET: **Efficient replication of a DNA base pair between non-hydrogen-bonded nucleoside analogues.** *Nature Struct Biol* 1998, **5**:950-954.

49. Guckian KM, Ren RXF, Chaudhuri NC, Tahmassebi DC, Kool ET: **Factors contributing to aromatic stacking in water: evaluation in the context of DNA.** *J Am Chem Soc* 2000, **122**:2213-2222.

50. Kool ET: **A mechanism for shape selection during DNA synthesis.** *Biopolymers (Nucleic Acid Sciences)* 1998, **48**:3-17.

51. Morales JC, Kool ET: **Varied molecular interactions in the active sites of several DNA polymerases: nonpolar nucleoside isosteres as probes.** *J Am Chem Soc* 2000, **122**:1001-1007.

52. Matray TJ, Kool ET: **A specific partner for abasic damage in DNA.** *Nature* 1999, **399**:704-708.

A highly efficient and selective replicable non-H-bonded base pair is described.

53. Matray TJ, Kool ET: **Selective and stable DNA base pairing without hydrogen bonds.** *J Am Chem Soc* 1998, **120**:6191-6192.

54. McMinn DL, Ogawa AK, Wu YQ, Liu JQ, Schultz PG, Romesberg FE: **Efforts toward expansion of the genetic alphabet: DNA polymerase recognition of a highly stable, self-pairing hydrophobic base.** *J Am Chem Soc* 1999, **121**:11585-11586.

A self-pairing, non-H-bonding base is described; it forms a strongly stable self-pair and is an efficient polymerase substrate.

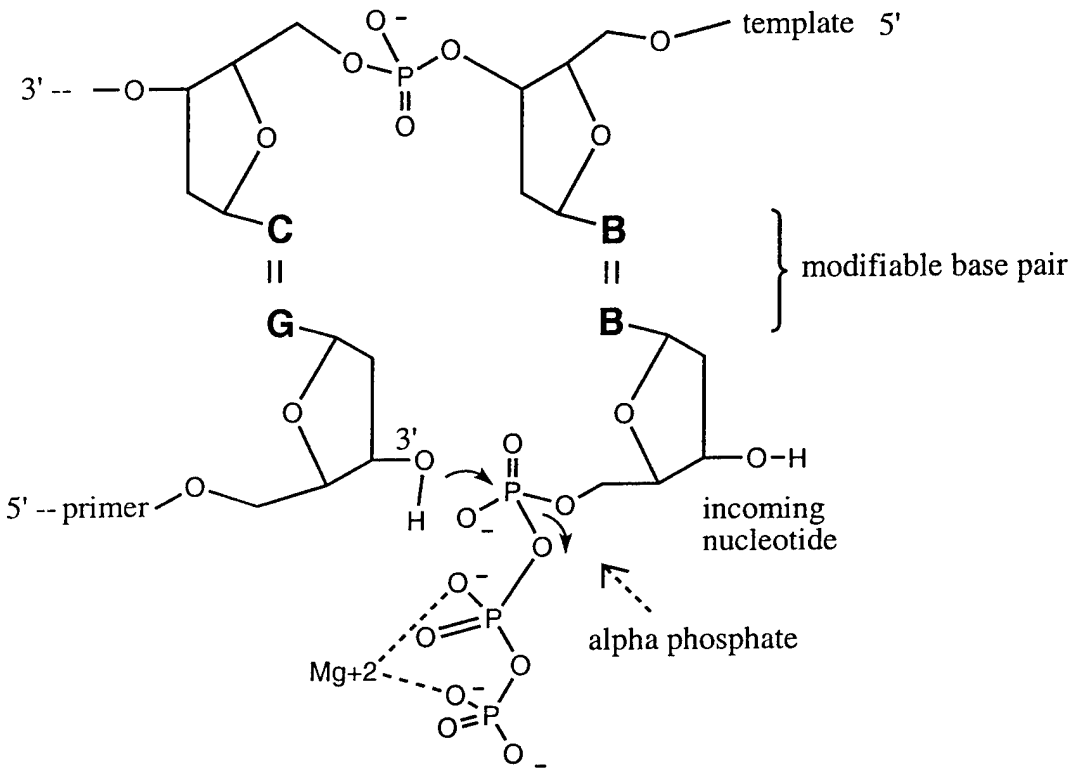
55. Ogawa AK, Wu YQ, McMinn DL, Liu JQ, Schultz PG, Romesberg FE: **Efforts toward the expansion of the genetic alphabet: information storage and replication with unnatural hydrophobic base pairs.** *J Am Chem Soc* 2000, **122**:3274-3287.

56. Guckian KM, Krugh TR, Kool ET: **Solution structure of a duplex DNA containing a replicable difluorotoluene-adenine pair**. *Nature Struct Biol* 1998, **5**:954-959.

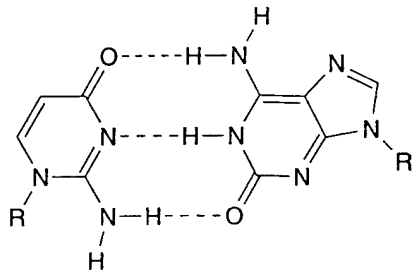
Figure legends

Figure 1. The chemistry of polymerase insertion of a nucleoside triphosphate. In the template, modifiable parts of the DNA include the phosphodiester linkage and the template base. For the incoming nucleotide, the alpha phosphate can be modified, as can the base.

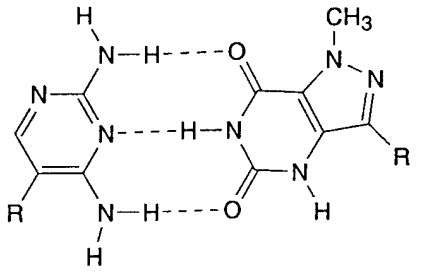
Figure 2. Structures of some functioning altered DNA base pairs. A. Altered H-bonding pairs of Benner. B. Altered functional pair of Rappaport. C. New base pair from the Yokoyama laboratory. D. Non-H-bonding pairs from Kool laboratory. E. Self-pair of Romesberg and Schultz.


template 5'
3' --
C
B
modifiable base pair
G
B
3'
O-H
5' -- primer
incoming
nucleotide
alpha phosphate
Mg+2

A

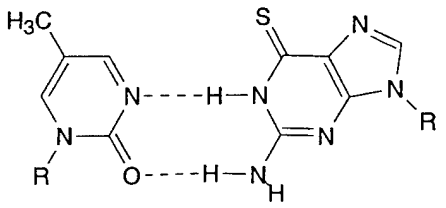


iso -C *iso* -G

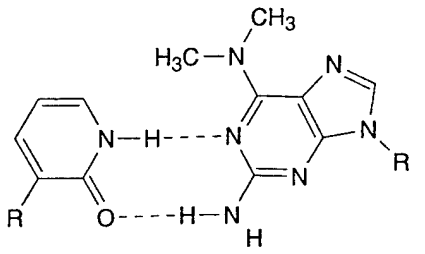


κ π

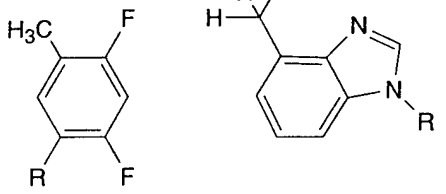
B



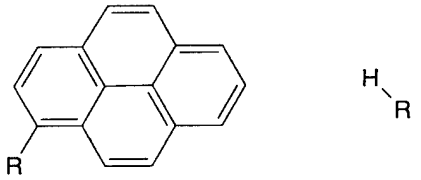
C



D



F Z



P ϕ

E